

EXTREME-STRIKE ASYMPTOTICS FOR GENERAL GAUSSIAN STOCHASTIC VOLATILITY MODELS

ARCHIL GULISASHVILI, FREDERI VIENS, AND XIN ZHANG

ABSTRACT. We consider a stochastic volatility stock price model in which the volatility is a non-centered continuous Gaussian process with arbitrary prescribed mean and covariance. By exhibiting a Karhunen-Loève expansion for the integrated variance, and using sharp estimates of the density of a general second-chaos variable, we derive asymptotics for the stock price density and implied volatility in these models in the limit of large or small strikes. Our main result provides explicit expressions for the first three terms in the expansion of the implied volatility, based on three basic spectral-type statistics of the Gaussian process: the top eigenvalue of its covariance operator, the multiplicity of this eigenvalue, and the L^2 norm of the projection of the mean function on the top eigenspace. Strategies for using this expansion for calibration purposes are discussed.

AMS 2010 Classification: 60G15, 91G20, 40E05.

Keywords: stochastic volatility, implied volatility, large strike, Karhunen-Loève expansion, chi-squared variates.

1. INTRODUCTION

In this article, we characterize the extreme-strike behavior of implied volatility curves for fixed maturity for uncorrelated Gaussian stochastic volatility models. This introduction contains a careful description of the problem's background and of our motivations. Before going into details, we summarize some of the article's specificities; all terminology in the next two paragraphs is referenced, defined, and/or illustrated in the remainder of this introduction.

We hold calibration of volatility smiles as a principal motivator. Cognizant of the fact that non-centered Gaussian volatility models can be designed in a flexible and parsimonious fashion, we adopt that class of models, imposing no further conditions on the marginal distribution of the volatility process itself, beyond pathwise continuity. The spectral structure of second-Wiener chaos variables allows us to work at that level of generality. We find that the first three terms in the extreme-strike implied volatility asymptotics – which is typically amply sufficient in applications – can be determined explicitly thanks to three parameters characterizing the top of the spectral decomposition of the integrated variance. In order to prove such a precise statement while relying on a moderate amount of technicalities, we make use of the simplifying assumption that the stochastic volatility is assumed independent of the stock price's driving noise.

When considering the trade-off between this restriction and calibration considerations, we observe that our model flexibility combined with known explicit spectral expansions and numerical tools may allow practitioners to compute the said spectral parameters in a straightforward fashion based on smile features, while also allowing them to select their favorite Gaussian volatility model class. Specific examples of Gaussian volatility processes are non-centered Brownian motion, Brownian bridge, and Ornstein-Uhlenbeck models. This last sub-class can be particularly appealing to practitioners since it contains stationary volatilities, and includes the well-known Stein-Stein model. We also mention how any Gaussian model specification, including long-memory ones, can be handled, thanks to the numerical ability to determine its spectral elements. We understand that the assumption of the stochastic volatility model being uncorrelated implies the symmetry of the implied volatility in the wings, which in some applications, is not a desirable feature; on the other hand, in many option markets, liquidity considerations limit the ability to calibrate using the large-strike wing (see the calibration study on SPX options in [20, Section 5.4]). The case of a correlated Gaussian stochastic volatility model is more complicated, but constitutes an interesting mathematical challenge, which we will investigate separately from this article, since one may need to develop completely new methods and techniques. An important step toward a better understanding of the asymptotic behavior of the implied volatility in some correlated stochastic volatility models is found in the articles [13, 14]. Another problem which is mathematically interesting and important in practice is the asymptotics for implied volatility in small or large time to maturity, on which we report in separate works.

1.1. Background and heuristics. Studies in quantitative finance based on the Black-Scholes-Merton framework have shown awareness of the inadequacy of the constant volatility assumption, particularly after the crash of 1987, when practitioners began considering that extreme events were more likely than what a log-normal model will predict. Propositions to exploit this weakness in log-normal modeling systematically and quantitatively have grown ubiquitous to the point that implied volatility (IV), or the volatility level that market call option prices would imply if the Black-Scholes model were underlying, is now a *bona fide* and vigorous topic of investigation, both at the theoretical and practical level. The initial evidence against constant volatility simply came from observing that IV as a function of strike prices for liquid call options exhibited non-constance, typically illustrated as a convex curve, often with a minimum near the money as for index options, hence the term ‘volatility smile’.

Stock price models where the volatility is a stochastic process are known as stochastic volatility models; the term ‘uncorrelated’ is added to refer to the submodel class in which the volatility process is independent of the noise term driving the stock price. In a sense, the existence of the smile for any uncorrelated stochastic volatility model was first proved mathematically by Renault and Touzi in [30]. They established that the IV as a function of the strike price decreases on the interval where the call is in the money, increases on the interval where the call is out of the money, and attains its minimum where the call is at the money. Note that Renault and Touzi did not prove that the IV is locally convex near the money, but their work still established stochastic volatility models as a main model class for studying IV; these models continued steadily to provide inspiration for IV studies.

A current emphasis, which has become fertile mathematical ground, is on IV asymptotics, such as large/small-strike, large-maturity, or small-time-to-maturity behaviors. These are helpful to understand and select models based on smile shapes. Several techniques are used to derive IV asymptotics. For instance, by exploiting a method of moments and the representation of power payoffs as mixtures of a continuum of calls with varying strikes, in a rather model-free context, R. Lee proved in [28] that, for models with positive moment explosions, the squared IV's large strike behavior is of order the log-moneyness $\log(K/S_0)$ times a constant which depends explicitly on supremum of the order of finite moments. A similar result holds for models with negative moment explosions, where the squared IV behaves like $K \mapsto \log(S_0/K)$ for small values of K . More general formulas describing the asymptotic behavior of the IV in the 'wings' ($K \rightarrow 0$ or $+\infty$) were obtained in [3, 4, 5, 22, 23, 25, 17] (see also the book [21]).

From the standpoint of modeling, one of the advantages of Lee's original result is the dependence of IV asymptotics merely on some simple statistics, namely as we mentioned, in the notation in [28], the maximal order \tilde{p} of finite moments for the underlying S_T , i.e.

$$\tilde{p}(T) := \sup \left\{ p \in \mathbb{R} : \mathbb{E} \left[(S_T)^{p+1} \right] < \infty \right\}.$$

This allows the author to draw appropriately strong conclusions about model calibration. A typical class in which \tilde{p} is positive and finite is that of Gaussian volatility models, which we introduce next.

We consider the stock price model of the following form:

$$dS_t = rS_t dt + |X_t|S_t dW_t : t \in [0, T], \quad (1)$$

where the short rate r is constant, $X(t) = m(t) + \tilde{X}(t)$ with m an arbitrary continuous deterministic function on $[0, T]$ (the mean function), and \tilde{X} is a continuous centered Gaussian process on $[0, T]$ independent of W , with arbitrary covariance Q . Note that it is not assumed in (36) that the process X is a solution to a stochastic differential equation as is often assumed in classical stochastic volatility models. A well-known special example of a Gaussian volatility model is the Stein-Stein model introduced in [33], in which the volatility process X is the so-called mean-reverting Ornstein-Uhlenbeck process satisfying

$$dX_t = \alpha(m - X_t) dt + \beta dZ_t \quad (2)$$

where m is the level of mean reversion, α is the mean-reversion rate, and β is level of uncertainty on the volatility; here Z is another Brownian motion, which may be correlated with W . In the present paper, we adopt an analytic technique, encountered for instance in the analysis of the uncorrelated Stein-Stein model by this paper's first author and E.M. Stein in [24] (see also [21]).

Returning to the question of the value of \tilde{p} , for a Gaussian volatility model, it can sometimes be determined by simple calculations, which we illustrate here with an elementary example. Assume S is a geometric Brownian motion with random volatility, i.e. a model as in (1) where (abusing notation) $|X_t|$ is taken the non-time-dependent $\sigma|X|$ where σ is a constant and X is an independent unit-variance normal variate (not dependent on t). Thus, at time T , with zero discount rate, $S_T = S_0 \exp(\sigma|X|W_T - \sigma^2 X^2 T/2)$. To simplify this example to the maximum, also assume that X is centered; using the independence of X and W , we get that we may replace $|X|$ by X in this example, since this does not change

the law of S_T (i.e. in the uncorrelated case, X 's non-positivity does not violate standard practice for volatility modeling). Then, using maturity $T = 1$, for any $p > 0$, the p th moment, via a simple change of variable, equals

$$\mathbb{E} [(S_1)^p] = \frac{1}{2\pi\sqrt{1+p\sigma^2}} \iint_{\mathbb{R}^2} dx dy \exp \left(-\frac{1}{2} \left(y^2 + w^2 - 2\frac{p\sigma}{\sqrt{1+p\sigma^2}} wy \right) \right)$$

which by an elementary computation is finite, and equal to $(1 + p\sigma^2) / (1 + p\sigma^2 - p^2\sigma^2)$, if and only if

$$p < \tilde{p} + 1 = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{1}{\sigma^2}}.$$

In the cases where the random volatility model X above is non-centered and is correlated with W , a similar calculation can be performed, at the essentially trivial expenses of invoking affine changes of variables, and the linear regression of one normal variate against another.

The above example illustrates heuristically that, by Lee's moment formula, the computation of \tilde{p} might be the quickest path to obtain the leading term in the large-strike expansion of the IV, for more complex Gaussian volatility models, namely ones where the volatility X is time-dependent. However, computing \tilde{p} is not necessarily an easy task, and appears, perhaps surprisingly, to have been performed rarely. For the Stein-Stein model, the value of \tilde{p} can be computed using the sharp asymptotic formulas for the stock price density near zero and infinity, established in [24] for the uncorrelated Stein-Stein model, and in [14] for the correlated one. These two papers also provide asymptotic formulas with error estimates for the IV at extreme strikes in the Stein-Stein model. Beyond the Stein-Stein model, little was known about the extreme strike asymptotics of general Gaussian stochastic volatility models. In the present paper, we extend the above-mentioned results from [24] and [14] to such models.

1.2. Motivation and summary of main result. Adopting the perspective that an asymptotic expansion for the IV can be helpful for model selection and calibration, our objective is to provide an expansion for the IV in a Gaussian volatility model relying on a minimal number of parameters, which can then be chosen to adjust to observed smiles. The restriction of non-correlated volatility means that the stock price distribution is a mixture model of geometric Brownian motions with time-dependent volatilities, whose mixing density at time T is that of the square root of a variable in the second-chaos of a Wiener process. That second-chaos variable is none other than the integrated variance

$$\Gamma_T := \int_0^T X_s^2 ds.$$

By relying on a general Hilbert-space structure theorem which applies to the second Wiener chaos, we prove that, in the most general case of a non-centered Gaussian stochastic volatility with a possible degeneracy in the eigenstructure of the covariance Q of X viewed as a linear operator on $L^2([0, T])$ (i.e. when the top eigenvalue λ_1 is allowed to have a multiplicity n_1 larger than 1), the large-strike IV asymptotics can be expressed with three terms which depend explicitly on T and on the following three parameters: λ_1 ,

n_1 , and the ratio

$$\delta = \|P_{E_1} m\|^2 / \lambda_1$$

where $\|P_{E_1} m\|$ is the norm in $L^2([0, T])$ of the orthogonal projection of the mean function m on the first eigenspace of Q . Specifically, with $I(K)$ the IV as a function of strike K , letting $k := \log(K/s_0) - rT$ be the discounted log-moneyness, as $k \rightarrow +\infty$, we prove

$$I(K) = L(T, \lambda_1) \sqrt{k} + M(T, \lambda_1, \delta) - \frac{n_1 - 1}{4} \frac{\log k}{\sqrt{k}} + O\left(\frac{1}{\sqrt{k}}\right) \quad (3)$$

where the constants L and M depend explicitly on T and λ_1 , and M also depends explicitly on δ . The details of these constants are in Theorem 4 on page 15. A similar asymptotic formula is obtained in the case where $k \rightarrow -\infty$, using symmetry properties of uncorrelated stochastic volatility models (see formula (50) on page 19). The specific case of the Stein-Stein model is expanded upon in some detail. Numerical illustrations of how the small strike asymptotics can be used are provided in Section 7, in the context of calibration; we explain some of the calibration ideas in the next subsection.

1.3. Practical implications. The first-order constant L is always strictly positive. The second-order term (the constant M) vanishes if and only if m is orthogonal to the first eigenspace of Q , which occurs for instance when $m \equiv 0$. The third-order term vanishes if and only if the top eigenvalue has multiplicity one, which is typical. The behavior of L and M as functions of T is determined partly by how the top eigenvalue λ_1 depends on T , which can be non-trivial. In the present paper, we assume T is fixed.

For fixed maturity T , assuming that Q has lead multiplicity $n_1 = 1$ for instance, a practitioner will have the possibility of determining a value λ_1 and a value δ to match the specific root-log-moneyness behavior of small- or large-strike IV; moreover in that case, choosing a constant mean function m , one obtains $\delta = m^2 \lambda_1^{-1} \left| \int_0^T e_1(t) dt \right|^2$ where e_1 is the top eigenfunction of Q . Market prices may not be sufficiently liquid at extreme strikes to distinguish between more than two parameters; this is typical of calibration techniques for implied volatility curves for fixed maturity, such as the ‘stochastic volatility inspired’ (SVI) parametrization disseminated by J. Gatheral: see [18, 19] (see also [20] and the references therein). Our result shows that Gaussian volatility models with non-zero mean are sufficient for this flexibility, and provide equivalent asymptotics irrespective of the precise mean function and covariance eigenstructure, since modulo the disappearance of the third-order term in the unit top multiplicity case $n_1 = 1$, only λ_1 and δ are relevant. Moreover, our Gaussian parametrization is free of arbitrage, since it is based on a semi-martingale model (1). In the case of SVI parametrization, the absence of arbitrage can be non-trivial, as discussed in [18] and [20].

Modelers wishing to stick to well-known classes of processes for X may then adjust the value of λ_1 by exploiting any available invariance properties for the desired class. For example, if X is standard Brownian motion, or the Brownian bridge, on $[0, T]$, we have $\lambda_1 = 4T^2/\pi$ or $\lambda_1 = T^2/\pi$ respectively, and these values scale quadratically with respect to a multiplicative scaling constant for X , beyond which an arbitrary mean value m may be chosen. If X is the mean-zero stationary OU process, we have $\lambda_1 = \beta^2 / (\omega_T + \alpha^2)$ where ω_T is the smallest positive solution of $2\alpha\omega \cos(\omega T) + (\alpha^2 - \omega) \sin(\omega T) = 0$, in

which case, for a fixed arbitrarily selected rate of mean reversion α , a scaling of λ_1 is then equivalent to selecting the variance of X , while a constant mean value m can then be selected independently. [9, Chapter 1] can be consulted for the eigenstructure of the covariance of Brownian motion and the Brownian bridge, which are classical results, and for a proof of the eigenstructure of the OU covariance (see also [11]). The top eigenfunctions in all three of these cases are known explicit trigonometric functions (see [9, Chapter 1]), and need to be referenced when selecting m . For the OU bridge, the eigenstructure of Q (equivalently known as the Karhunen-Loève expansion of Q) was found in [10], while in [12], such an expansion was characterized for special Gaussian processes generated by independent pairs of exponential random variables. On the other hand, fractional Brownian motion and OU processes driven by fractional Brownian motion do not fall in the class of Gaussian processes for which the Karhunen-Loève expansion is known explicitly. However, efficient numerical techniques exist to compute the eigenfunctions and eigenvalues in these cases: see [9, Chapter 2]. OU processes driven by fractional Brownian motion were proposed early on for option pricing, and recently analyzed in [8, 7]. Section 7 illustrates how, in the case of the Stein-Stein model and of its version driven by fraction Brownian motion, the explicit, semi-explicit, or numerically accessible Karhunen-Loève expansion of X can be used in conjunction with the asymptotics (3) for calibrating parameters. We find that market liquidity considerations limit the theoretical range of applicability of calibration strategies, but that significant practical results are nonetheless available.

The remainder of this article is structured as follows. Section 2 sets up a convenient second-chaos representation for the model's integrated volatility. Section 3 uses calculations in a proof in [6] to derive precise asymptotics for the density of the mixing distribution. Section 4 converts these asymptotics into two-sided estimates of the density of the stock price S_T , thanks to the analytic tools developed in [24, 21]. Sharper asymptotic formulas for the density can also be obtained, but they are not needed to derive sharp asymptotic formulas for the IV with three terms and an error estimate. In Section 5, we characterize the wing behavior of the implied volatility in Gaussian stochastic volatility models. The special case of the uncorrelated Stein-Stein model is studied in more detail in Section 6. Finally, our practical study of calibration strategies, with numerics, is in Section 7.

2. GENERAL SETUP AND SECOND-CHAOS EXPANSION OF THE INTEGRATED VARIANCE

Let X be an almost-surely continuous Gaussian process on a filtered complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ with mean and covariance functions denoted by $m(t) = \mathbb{E}[X_t]$ and

$$Q(t, s) = \text{cov}(X_t, X_s) = \mathbb{E}[(X_t - m(t))(X_s - m(s))],$$

respectively. While such processes used in a jump-free quantitative finance context for volatility modeling will require, in addition, that X be adapted to filtration of the Wiener process W driving the stock-price (as in (1)), under our simplifying assumption that X and W be independent, this adaptability assumption can be considered as automatically satisfied, or equivalently, as unnecessary, since the filtration of W can be augmented by

the natural filtration of X . Define the centered version of X :

$$\tilde{X}_t := X_t - m(t), \quad t \geq 0.$$

Fix a time horizon $T > 0$. It is not hard to see that $Q(s, s) > 0$ for all $s > 0$. Since the Gaussian process X is almost surely continuous, the mean function $t \mapsto m(t)$ is a continuous function on $[0, T]$, and the covariance function $(t, s) \mapsto Q(t, s)$ is a continuous function of two variables on $[0, T]^2$. This is a consequence of the Dudley-Fernique theory of regularity, which also implies that m and Q boast moduli of continuity bounded above by the scale $h \mapsto \log^{-1/2}(h^{-1})$ (see [1]), but this can also be established by elementary means¹.

In our analysis, it will be convenient to refer to the Karhunen-Loève expansion of \tilde{X} . Applying the classical Karhunen-Loève theorem to \tilde{X} (see, e.g., [34], Section 26.1), we obtain the existence of a non-increasing sequence of non-negative summable reals $\{\lambda_n : n = 1, 2, \dots\}$, an i.i.d. sequence of standard normal variates $\{Z_n : n = 1, 2, \dots\}$, and a sequence of functions $\{e_n : n = 1, 2, \dots\}$ which form an orthonormal system in $L^2([0, T])$, such that

$$\tilde{X}_t = \sum_{n=1}^{\infty} \sqrt{\lambda_n} e_n(t) Z_n. \quad (4)$$

In (4), $\{e_n = e_{n,T}\}$ are the eigenfunctions of the covariance Q acting on $L^2([0, T])$ as the following operator

$$\mathcal{K}(f)(t) = \int_0^T f(s) Q(t, s) ds, \quad f \in L^2([0, T]), \quad 0 \leq t \leq T,$$

and $\{\lambda_n = \lambda_{n,T}\}$, $n \geq 1$, are the corresponding eigenvalues (counting the multiplicities). We always assume that the orthonormal system $\{e_n\}$ is rearranged so that

$$\lambda_1 = \lambda_2 = \dots = \lambda_{n_1} > \lambda_{n_1+1} = \lambda_{n_1+2} = \dots = \lambda_{n_1+n_2} > \dots$$

In particular, λ_1 is the top eigenvalue, and n_1 is its multiplicity.

Using (4), we obtain

$$\int_0^T \tilde{X}_t^2 dt = \int_0^T \left(\sum_{n=1}^{\infty} \sqrt{\lambda_n} e_n(t) Z_n \right)^2 dt = \sum_{n=1}^{\infty} \lambda_n Z_n^2. \quad (5)$$

It is worth pointing out that this expression for the integrated variance of the centered volatility $\int_0^T \tilde{X}_t^2 dt$ is in fact the most general form of a random variable in the second Wiener chaos, with mean adjusted to ensure almost-sure positivity of the integrated variance. This is established using a classical structure theorem on separable Hilbert spaces, as explained in [29, Section 2.7.4]. In other words (also see [29, Section 2.7.3] for additional

¹The continuity of the process X implies its continuity in probability on Ω . Hence, the process X is continuous in the mean-square sense (see, e.g., [27], Lemma 1 on p. 5, or invoke the equivalence of L^p norms on Wiener chaos, see [29]). Mean-square continuity of X implies the continuity of the mean function on $[0, T]$. In addition, the autocorrelation function of the process X , that is, the function $R(t, s) = \mathbb{E}[X_t X_s]$, $(t, s) \in [0, T]^2$, is continuous (see, e.g., [2], Lemma 4.2). Finally, since $Q(t, s) = R(t, s) - m(t)m(s)$, the covariance function Q is continuous on $[0, T]^2$.

details), any prescribed mean-adjusted integrated variance in the second chaos is of the form

$$V(T) := \iint_{[0,T]^2} G(s,t) dZ(s) dZ(t) + 2 \|G\|_{L^2([0,T]^2)}^2$$

for some standard Wiener process Z and some function $G \in L^2([0,T]^2)$, and moreover one can find a centered Gaussian process \tilde{X} such that $V(T) = \int_0^T \tilde{X}_t^2 dt$ and one can compute the coefficients λ_n in the Karhunen-Loève representation 5 as the eigenvalues of the covariance of \tilde{X} . When using the non-centered process X , this analysis immediately yields

$$\begin{aligned} \int_0^T X_t^2 dt &= \int_0^T (\tilde{X}_t + m(t))^2 dt \\ &= \sum_{n=1}^{\infty} \lambda_n Z_n^2 + 2 \sum_{n=1}^{\infty} \sqrt{\lambda_n} \left[\int_0^T m(t) e_n(t) dt \right] Z_n \\ &\quad + \int_0^T m(t)^2 dt. \end{aligned} \tag{6}$$

Set

$$\delta_n = \delta_{n,T} = \int_0^T m(t) e_n(t) dt, \quad n \geq 1, \tag{7}$$

and

$$s = s(T) = \int_0^T m(t)^2 dt. \tag{8}$$

Then Bessel's inequality implies that

$$\sum_{n=1}^{\infty} \delta_n^2 \leq s. \tag{9}$$

Denote

$$\tau = s - \sum_{n=1}^{\infty} \delta_n^2. \tag{10}$$

It is not hard to see that if the function $t \mapsto m(t)$ belongs to the image space $\mathcal{K}(L^2[0,T])$, then

$$\sum_{n=1}^{\infty} \delta_n^2 = s, \tag{11}$$

and hence, $\tau = 0$. For instance, the previous equality holds for a centered Gaussian process X .

Equality (6) can be rewritten as follows:

$$\begin{aligned} \int_0^T X_t^2 dt &= \sum_{n=1}^{\infty} \lambda_n \left[Z_n^2 + 2 \frac{\delta_n}{\sqrt{\lambda_n}} Z_n \right] + s \\ &= \sum_{n=1}^{\infty} \lambda_n \left[Z_n + \frac{\delta_n}{\sqrt{\lambda_n}} \right]^2 + \left(s - \sum_{n=1}^{\infty} \delta_n^2 \right). \end{aligned} \tag{12}$$

It follows from (11) and (12) that if the function $t \mapsto m(t)$ belongs to the image space $\mathcal{K}(L^2[0, T])$, then

$$\int_0^T X_t^2 dt = \sum_{n=1}^{\infty} \lambda_n \left[Z_n + \frac{\delta_n}{\sqrt{\lambda_n}} \right]^2. \quad (13)$$

Let us denote the noncentral chi-square distribution with the number of degrees of freedom k and the parameter of noncentrality λ by $\chi^2(k, \lambda)$ (more information on such distributions can be found in [21] or in any probability textbook). Define a random variable \tilde{Z}_n by

$$\tilde{Z}_n = \left[Z_n + \frac{\delta_n}{\sqrt{\lambda_n}} \right]^2.$$

It is clear that \tilde{Z}_n is distributed as $\chi^2(1, \frac{\delta_n^2}{\lambda_n})$. Set

$$\rho_1 = \lambda_1, \quad \rho_2 = \lambda_{n_1+1}, \quad \rho_3 = \lambda_{n_1+n_2+1}, \quad \dots,$$

and

$$\Lambda_T = \frac{1}{\lambda_1} \left(\int_0^T X_t^2 dt - \tau \right). \quad (14)$$

Then, using (12), we see that

$$\Lambda_T = \chi^2 \left(n_1, \frac{1}{\lambda_1} \sum_{n=1}^{n_1} \delta_n^2 \right) + \frac{\rho_2}{\lambda_1} \chi^2 \left(n_2, \frac{1}{\rho_2} \sum_{n=n_1+1}^{n_1+n_2} \delta_n^2 \right) + \dots, \quad (15)$$

where the repeated chi-squared notation is used abusively to denote independent chi-squared random variables.

3. ASYMPTOTICS OF THE MIXING DENSITY

The asymptotic behavior of the complementary distribution function of an infinite linear combination of independent central chi-square random variables was characterized by Zolotarev in [35]. Zolotarev's results were generalized to the case of noncentral chi-square variables by Beran (see [6]). Beran used some ideas from [26] in his work. We will employ Beran's results in the present paper. More precisely, Theorem 2 in [6] will be used below. This theorem provides an asymptotic formula for the complementary distribution function of an infinite linear combination of independent noncentral chi-square random variables. A sharper formula for the distribution density q_T of such a random variable can be extracted from the proof of Theorem 2 in [6] (see the very end of that proof). Adapting Beran's result to our case and taking into account estimate (9) (this estimate is needed to check the validity of the conditions in Beran's theorem), we see that

$$\left| \frac{q_T(x)}{p_{\chi^2} \left(x; n_1, \frac{1}{\lambda_1} \sum_{n=1}^{n_1} \delta_n^2 \right)} - A \right| = O \left(x^{-\frac{1}{2}} \right) \quad (16)$$

as $x \rightarrow \infty$. In (16), the number A is given by

$$A = \prod_{k=2}^{\infty} \left(\frac{\lambda_1}{\lambda_1 - \rho_k} \right)^{\frac{n_k}{2}} \times \exp \left\{ \frac{1}{2} \sum_{k=2}^{\infty} \frac{1}{\lambda_1 - \rho_k} \left(\sum_{n=n_1+\dots+n_{k-1}+1}^{n_1+\dots+n_k} \delta_n^2 \right) \right\}. \quad (17)$$

Set

$$\delta = \frac{1}{\lambda_1} \sum_{n=1}^{n_1} \delta_n^2. \quad (18)$$

Then (16) gives

$$q_T(x) = A p_{\chi^2}(x; n_1, \delta) \left(1 + O\left(x^{-\frac{1}{2}}\right) \right). \quad (19)$$

For $\lambda > 0$, the following formula is known:

$$p_{\chi^2}(x; n, \lambda) = \frac{1}{2} \left(\frac{x}{\lambda} \right)^{\frac{n}{4} - \frac{1}{2}} e^{-\frac{x+\lambda}{2}} I_{\frac{n}{2}-1}(\sqrt{\lambda x}), \quad x > 0, \quad (20)$$

where I_ν is the modified Bessel function of the first kind (see, e.g., [21], Theorem 1.31). It is easy to see that formula (20) and the formula

$$I_\nu(t) = \frac{e^t}{\sqrt{2\pi t}} \left(1 + O\left(t^{-1}\right) \right), \quad t \rightarrow \infty,$$

describing the asymptotic behavior of the I-Bessel function, imply that

$$p_{\chi^2}(x; n, \lambda) = \frac{1}{2\sqrt{2\pi}} \lambda^{-\frac{n-1}{4}} x^{\frac{n-3}{4}} e^{\sqrt{\lambda x}} e^{-\frac{x+\lambda}{2}} \left(1 + O\left(x^{-\frac{1}{2}}\right) \right) \quad (21)$$

as $x \rightarrow \infty$. On the other hand, if $\lambda = 0$, then

$$p_{\chi^2}(x; n, 0) = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n-2}{2}} \exp\left\{-\frac{x}{2}\right\}, \quad x > 0, \quad (22)$$

(see, e.g., Lemma 1.27 in [21]).

Recall that we denoted by q_T the distribution density of the random variable Λ_T defined by (14). It follows from (19) and (21) that

$$q_T(x) = \frac{A}{2\sqrt{2\pi}} \delta^{-\frac{n_1-1}{4}} x^{\frac{n_1-3}{4}} e^{\sqrt{\delta x}} e^{-\frac{x+\delta}{2}} \left(1 + O\left(x^{-\frac{1}{2}}\right) \right) \quad (23)$$

as $x \rightarrow \infty$. The constants A and δ in (23) are defined by (17) and (18), respectively. For a centered Gaussian process X , (19) and (22) imply that

$$q_T(x) = \frac{A}{2^{\frac{n_1}{2}} \Gamma\left(\frac{n_1}{2}\right)} x^{\frac{n_1-2}{2}} \exp\left\{-\frac{x}{2}\right\} \left(1 + O\left(x^{-\frac{1}{2}}\right) \right), \quad x > 0, \quad (24)$$

as $x \rightarrow \infty$, where

$$A = \prod_{k=2}^{\infty} \left(\frac{\lambda_1}{\lambda_1 - \rho_k} \right)^{\frac{n_k}{2}}. \quad (25)$$

Indeed, in this case, we have $s = 0$, $\delta_n = 0$ for all $n \geq 1$, $\delta = 0$, and $\tau = 0$.

Our main goal is to characterize the asymptotic behavior of the distribution density p_T of the random variable

$$\Gamma_T = \int_0^T X_t^2 dt. \quad (26)$$

It follows from (10) and (14) that $\Gamma_T = \lambda_1 \Lambda_T + \tau$. Therefore,

$$p_T(x) = \frac{1}{\lambda_1} q_T \left(\frac{1}{\lambda_1} (x - \tau) \right). \quad (27)$$

Theorem 1. *Let p_T be the distribution density of the random variable Γ_T given by (26). If $\delta > 0$, then the following asymptotic formula holds:*

$$\begin{aligned} p_T(x) &= C x^{\frac{n_1-3}{4}} \exp \left\{ \sqrt{\frac{\delta}{\lambda_1}} \sqrt{x} \right\} \exp \left\{ -\frac{x}{2\lambda_1} \right\} \\ &\quad \times \left(1 + O \left(x^{-\frac{1}{2}} \right) \right) \end{aligned} \quad (28)$$

as $x \rightarrow \infty$, where

$$C = \frac{A}{2\sqrt{2\pi}} \lambda_1^{-\frac{1}{2}} \left(\sum_{n=1}^{n_1} \delta_n^2 \right)^{-\frac{n_1-1}{4}} \exp \left\{ \frac{s - \sum_{n=1}^{\infty} \delta_n^2 - \sum_{n=1}^{n_1} \delta_n^2}{2\lambda_1} \right\}. \quad (29)$$

In (28) and (29), the constants δ_n , s , A , and δ are defined by (7), (8), (17), and (18), respectively. On the other hand, for a centered Gaussian process X , we have

$$p_T(x) = C x^{\frac{n_1-2}{2}} \exp \left\{ -\frac{x}{2\lambda_1} \right\} \left(1 + O \left(x^{-\frac{1}{2}} \right) \right) \quad (30)$$

as $x \rightarrow \infty$, where

$$C = \frac{A}{2^{\frac{n_1}{2}} \Gamma \left(\frac{n_1}{2} \right)} \lambda_1^{-\frac{n_1}{2}}. \quad (31)$$

Proof. Formulas (23) and (27) imply that

$$\begin{aligned} p_T(x) &= \frac{A}{2\sqrt{2\pi}} \frac{1}{\lambda_1} \lambda_1^{\frac{n_1-1}{4}} \delta^{-\frac{n_1-1}{4}} \\ &\quad \times \left(\sum_{n=1}^{n_1} \delta_n^2 \right)^{-\frac{n_1-1}{4}} \lambda_1^{-\frac{n_1-3}{4}} \exp \left\{ \frac{\tau - \sum_{n=1}^{n_1} \delta_n^2}{2\lambda_1} \right\} \\ &\quad \times (x - \tau)^{\frac{n_1-3}{4}} \exp \left\{ \sqrt{\frac{\delta(x - \tau)}{\lambda_1}} \right\} \exp \left\{ -\frac{x}{2\lambda_1} \right\} \\ &\quad \times \left(1 + O \left(x^{-\frac{1}{2}} \right) \right) \end{aligned} \quad (32)$$

as $x \rightarrow \infty$.

Next, taking into account the formulas

$$(x - \tau)^{\frac{n_1-3}{4}} = x^{\frac{n_1-3}{4}} (1 + O(x^{-1}))$$

and

$$\exp \left\{ \sqrt{\frac{\delta(x-\tau)}{\lambda_1}} \right\} = \exp \left\{ \sqrt{\frac{\delta}{\lambda_1}} \sqrt{x} \right\} (1 + O(x^{-\frac{1}{2}})),$$

and simplifying the expression on the right-hand side of (32), we obtain formula (28). The proof of formula (30) is similar. Here we use (24) and (27). \square

The next assertion follows from Theorem 1.

Corollary 2. *Let p_T be the distribution density of the random variable Γ_T given by (26). Then the following are true:*

(1) *If $n_1 = 1$, then*

$$\begin{aligned} p_T(x) = & Cx^{-\frac{1}{2}} \exp \left\{ \sqrt{\frac{\delta}{\lambda_1}} \sqrt{x} \right\} \exp \left\{ -\frac{x}{2\lambda_1} \right\} \\ & \times \left(1 + O\left(x^{-\frac{1}{2}}\right) \right) \end{aligned} \quad (33)$$

as $x \rightarrow \infty$.

(2) *Suppose X is a centered Gaussian process with $n_1 = 1$. Then*

$$p_T(x) = Cx^{-\frac{1}{2}} \exp \left\{ -\frac{x}{2\lambda_1} \right\} \left(1 + O\left(x^{-\frac{1}{2}}\right) \right) \quad (34)$$

as $x \rightarrow \infty$.

4. STOCK PRICE ASYMPTOTICS

In this section, we study stochastic volatility models, for which the volatility is described by the absolute value of a Gaussian process.

Recall that in the present paper we assume that the asset price process S satisfies the following linear stochastic differential equation:

$$dS_t = rS_t dt + |X_t| S_t dW_t, \quad (35)$$

where X is a continuous Gaussian process on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$. In (35), W is a standard Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to the filtration $\{\mathcal{F}_t\}$, and the symbol r stands for the constant interest rate. It will be assumed that the processes X and W are independent. The initial price of the asset will be denoted by s_0 .

Since (35) is a linear stochastic differential equation, we have

$$S_t = s_0 \exp \left\{ rt - \frac{1}{2} \int_0^t X_s^2 ds + \int_0^t |X_s| dW_s \right\}.$$

The previous equality follows from the Doléans-Dade formula (see [31]). Therefore, the discounted asset price process is given by the following stochastic exponential:

$$\tilde{S}_t = e^{-rt} S_t = s_0 \exp \left\{ -\frac{1}{2} \int_0^t X_s^2 ds + \int_0^t |X_s| dW_s \right\}. \quad (36)$$

The next assertion states that under the restrictions on the volatility process used in the present paper, we are in a risk-neutral environment.

Lemma 3. *Under the restrictions on the volatility process X in (35), the discounted asset price process \tilde{S} is a $\{\mathcal{F}_t\}$ -martingale.*

Proof. Fix a time horizon $T > 0$. It suffices to prove that there exists $\delta > 0$ such that

$$L = \sup_{0 < t \leq T} \mathbb{E}[\exp\{\delta X_t^2\}] < \infty \quad (37)$$

(see, e.g., [21], Corollary 2.11). It will be shown that (37) holds provided that

$$\delta < \frac{1}{2 \max_{0 \leq t \leq T} Q(t, t)}. \quad (38)$$

We have

$$\begin{aligned} L &= \frac{1}{\sqrt{2\pi}} \sup_{0 < t \leq T} \frac{1}{\sqrt{Q(t, t)}} \int_{\mathbb{R}} \exp \left\{ \delta y^2 - \frac{1}{2Q(t, t)} (y - m(t))^2 \right\} \\ &\leq \frac{1}{\sqrt{2\pi}} \sup_{0 < t \leq T} \frac{\exp \left\{ -\frac{m(t)^2}{2Q(t, t)} \right\}}{\sqrt{Q(t, t)}} \int_{\mathbb{R}} \exp \left\{ -\left(\frac{1}{2Q(t, t)} - \delta \right) y^2 + \frac{m(t)}{Q(t, t)} y \right\} dy. \end{aligned}$$

Set $\alpha(t) = \frac{1}{2Q(t, t)} - \delta$ and $\beta(t) = \frac{m(t)}{Q(t, t)}$. Then, completing the square in the previous exponential, we obtain

$$\begin{aligned} L &\leq \frac{1}{\sqrt{2\pi}} \sup_{0 < t \leq T} \frac{\exp \left\{ -\frac{m(t)^2}{2Q(t, t)} \right\}}{\sqrt{Q(t, t)}} \exp \left\{ \frac{\beta(t)^2}{4\alpha(t)} \right\} \\ &\quad \int_{\mathbb{R}} \exp \left\{ -\alpha(t) \left(y - \frac{\beta(t)}{2\alpha(t)} \right)^2 \right\} dy \\ &= \frac{1}{\sqrt{2\pi}} \sup_{0 < t \leq T} \frac{1}{\sqrt{Q(t, t)}} \exp \left\{ \frac{\beta(t)^2}{4\alpha(t)} - \frac{m(t)^2}{2Q(t, t)} \right\} \int_{\mathbb{R}} \exp \left\{ -\alpha(t) y^2 \right\} dy \\ &= a_1 \sup_{0 < t \leq T} \frac{1}{\sqrt{Q(t, t)\alpha(t)}} \exp \left\{ \frac{\beta(t)^2}{4\alpha(t)} - \frac{m(t)^2}{2Q(t, t)} \right\} \\ &= a_2 \sup_{0 < t \leq T} \frac{1}{\sqrt{1 - 2\delta Q(t, t)}} \exp \left\{ \frac{m(t)^2}{2Q(t, t)} \left[\frac{1}{1 - 2\delta Q(t, t)} - 1 \right] \right\} \\ &= a_2 \sup_{0 < t \leq T} \frac{1}{\sqrt{1 - 2\delta Q(t, t)}} \exp \left\{ \frac{\delta m(t)^2}{1 - 2\delta Q(t, t)} \right\} \\ &\leq a_2 \frac{1}{\sqrt{1 - 2\delta \max_{0 \leq t \leq T} Q(t, t)}} \exp \left\{ \frac{\delta m(t)^2}{1 - 2\delta \max_{0 \leq t \leq T} Q(t, t)} \right\}. \end{aligned}$$

Finally, using (38) and the estimate $|m(t)| < M$ for all $0 \leq t \leq T$, we obtain (37).

This completes the proof of Lemma 3. \square

Since the processes X and W are independent, the following formula holds for the distribution density D_t of the asset price S_t :

$$D_t(x) = \frac{\sqrt{s_0 e^{rt}}}{\sqrt{2\pi t}} x^{-\frac{3}{2}} \int_0^\infty y^{-1} \exp \left\{ - \left[\frac{\log^2 \frac{x}{s_0 e^{rt}}}{2ty^2} + \frac{ty^2}{8} \right] \right\} \tilde{p}_t(y) dy. \quad (39)$$

In (39), \tilde{p}_t is the distribution density of the random variable

$$\tilde{Y}_t = \left\{ \frac{1}{t} \int_0^t X_s^2 ds \right\}^{\frac{1}{2}}.$$

The function \tilde{p}_t is called the mixing density. The proof of formula (39) can be found in [21] (see (3.5) in [21]).

It is not hard to see that

$$\tilde{p}_t(y) = 2ty p_t(ty^2), \quad (40)$$

where the symbol p_t stands for the density of the realized volatility $Y_t = \int_0^t X_s^2 ds$. It follows from formula (28) that

$$\begin{aligned} \tilde{p}_t(y) &= \tilde{A} y^{\frac{n_1-1}{2}} \exp \left\{ \tilde{B} y \right\} \exp \left\{ -\tilde{C} y^2 \right\} \\ &\times \left(1 + O(y^{-1}) \right) \end{aligned} \quad (41)$$

as $y \rightarrow \infty$, where

$$\tilde{A} = 2Ct^{\frac{n_1+1}{4}}, \quad \tilde{B} = \sqrt{\frac{\delta t}{\lambda_1}}, \quad \tilde{C} = \frac{t}{2\lambda_1}. \quad (42)$$

Our next goal is to estimate the function D_t . The asymptotic behavior as $x \rightarrow \infty$ of the integral appearing in (39) was studied in [24] (see also Section 5.3 in [24]). It is explained in [24] how to get an asymptotic formula for the integral in (39) in the case where an asymptotic formula for the mixing density is similar to formula (41). We refer the reader to the derivation of Theorem 6.1 in [21], which is based on formula (5.133) in Section 5.6 of [21] and Theorem 5.5 in [21]. The latter theorem concerns the asymptotic behavior of integrals with lognormal kernels. Having obtained an asymptotic formula for the distribution density of the asset price, we can find a similar asymptotic formula for the call pricing function C at large strikes, and then obtain an asymptotic formula for the implied volatility I (see Section 10.5 in [21]). However, the coefficients in the fourth term (and the higher order terms) in the asymptotic expansion of the implied volatility I are complicated and not suitable for computations. For that reason, we will restrict ourselves to asymptotic formulas for the implied volatility with three terms and an error estimate. To obtain such formulas, it is not necessary to use sharp asymptotic for D_t or C . It suffices to establish two-sided estimates for D_t and C (see [21], Chapters 9 and 10). The discussion below exploits the following two-sided estimate for the mixing density \tilde{p}_t (this estimate follows from (41):

$$\tilde{p}_t(y) \approx y^{\frac{n_1-1}{2}} \exp \left\{ \tilde{B} y \right\} \exp \left\{ -\tilde{C} y^2 \right\} \quad (43)$$

as $y \rightarrow \infty$. Using formula (39) and reasoning as was described above, we see that for $T > 0$,

$$D_T(x) \approx \left(\log \frac{x}{s_0 e^{rt}} \right)^{\frac{n_1-3}{4}} \exp \left\{ \frac{\tilde{B}\sqrt{2}}{T^{\frac{1}{4}}(8\tilde{C}+T)^{\frac{1}{4}}} \sqrt{\log \frac{x}{s_0 e^{rt}}} \right\} x^{-\left(\frac{3}{2} + \frac{\sqrt{8\tilde{C}+T}}{2\sqrt{T}}\right)} \quad (44)$$

as $x \rightarrow \infty$. In (44),

$$\tilde{B} = \sqrt{\frac{\delta T}{\lambda_1}} \quad \text{and} \quad \tilde{C} = \frac{T}{2\lambda_1}. \quad (45)$$

The proof of (44) uses (43) and is similar to the proof of formula (6.49) on pages 186-187 of [21].

5. ASYMPTOTICS OF THE IMPLIED VOLATILITY

The call pricing function in the stochastic volatility model described by (35) will be denoted by C . We have

$$C(T, K) = e^{-rt} \mathbb{E} [(S_T - K)^+]$$

where T is the maturity and K is the strike price. We will fix T , and consider C as the function $K \mapsto C(K)$ of only the strike price K . The Black-Scholes implied volatility associated with the pricing function C will be denoted by I . More information on the implied volatility can be found in [19, 21].

The next assertion provides an asymptotic formula for the implied volatility in the stochastic volatility model given by (35).

Theorem 4. *The following formula holds for the implied volatility $K \mapsto I(K)$ as $K \rightarrow \infty$:*

$$\begin{aligned} I(K) = & \frac{1}{T^{\frac{3}{4}}} \left[\left(\sqrt{8\tilde{C}+T} + \sqrt{T} \right)^{\frac{1}{2}} - \left(\sqrt{8\tilde{C}+T} - \sqrt{T} \right)^{\frac{1}{2}} \right] \sqrt{\log \frac{K}{s_0 e^{rT}}} \\ & + \frac{\sqrt{2}\tilde{B}}{T^{\frac{1}{2}}(8\tilde{C}+T)^{\frac{1}{4}}} \left[\frac{1}{\left(\sqrt{8\tilde{C}+T} - \sqrt{T} \right)^{\frac{1}{2}}} - \frac{1}{\left(\sqrt{8\tilde{C}+T} + \sqrt{T} \right)^{\frac{1}{2}}} \right] \\ & + \frac{1-n_1}{4} \frac{\log \log \frac{K}{s_0 e^{rT}}}{\sqrt{\log \frac{K}{s_0 e^{rT}}}} + O \left(\left(\log \frac{K}{s_0 e^{rT}} \right)^{-\frac{1}{2}} \right). \end{aligned} \quad (46)$$

The constants in (46) are defined by

$$\tilde{B} = \sqrt{\frac{T}{(\lambda_1)^2} \sum_{n=1}^{n_1} \delta_n^2} \quad \text{and} \quad \tilde{C} = \frac{T}{2\lambda_1}. \quad (47)$$

Proof. The equalities in (47) follow from (18) and (45).

We will use the following assertion which was established in [21] (see Theorem 10.17 in [21]).

Theorem 5. Suppose that the density of the asset price D_T satisfies

$$D_T(x) \approx x^\alpha h(x)$$

as $x \rightarrow \infty$, where $\alpha < -2$ and h is a slowly varying function. Then

$$\begin{aligned} I(K) &= \frac{\sqrt{2}}{\sqrt{T}} \sqrt{\log K + \log \frac{1}{K^{\alpha+2}h(K)} - \frac{1}{2} \log \log \frac{1}{K^{\alpha+2}h(K)}} \\ &\quad - \frac{\sqrt{2}}{\sqrt{T}} \sqrt{\log \frac{1}{K^{\alpha+2}h(K)} - \frac{1}{2} \log \log \frac{1}{K^{\alpha+2}h(K)}} \\ &\quad + O\left((\log K)^{-\frac{1}{2}}\right) \end{aligned}$$

as $K \rightarrow \infty$.

Set $\alpha = -\left(\frac{3}{2} + \frac{\sqrt{8\tilde{C}+T}}{2\sqrt{T}}\right)$ and

$$h(x) = \left(\log \frac{x}{s_0 e^{rT}}\right)^{\frac{n_1-3}{4}} \exp \left\{ \frac{\tilde{B}\sqrt{2}}{T^{\frac{1}{4}}(8\tilde{C}+T)^{\frac{1}{4}}} \sqrt{\log \frac{x}{s_0 e^{rT}}} \right\}.$$

It is easy to see that $\alpha < -2$ and h is a slowly varying function. Next, taking into account (44) and applying Theorem 5, we obtain

$$\begin{aligned} I(K) &= \frac{\sqrt{2}}{\sqrt{T}} \sqrt{(-\alpha-1) \log \frac{K}{s_0 e^{rT}} + \log \frac{1}{h(K)} - \frac{1}{2} \log \log \frac{1}{K^{\alpha+2}h(K)}} \\ &\quad - \frac{\sqrt{2}}{\sqrt{T}} \sqrt{(-\alpha-2) \log \frac{K}{s_0 e^{rT}} + \log \frac{1}{h(K)} - \frac{1}{2} \log \log \frac{1}{K^{\alpha+2}h(K)}} \\ &\quad + O\left(\left(\log \frac{K}{s_0 e^{rT}}\right)^{-\frac{1}{2}}\right) \end{aligned}$$

as $K \rightarrow \infty$. Now, using the mean value theorem, we get

$$\begin{aligned}
I(K) &= \frac{\sqrt{2}}{\sqrt{T}} \sqrt{(-\alpha - 1) \log \frac{K}{s_0 e^{rT}} + \log \frac{1}{h(K)} - \frac{1}{2} \log \log \frac{K}{s_0 e^{rT}}} \\
&\quad - \frac{\sqrt{2}}{\sqrt{T}} \sqrt{(-\alpha - 2) \log \frac{K}{s_0 e^{rT}} + \log \frac{1}{h(K)} - \frac{1}{2} \log \log \frac{K}{s_0 e^{rT}}} \\
&\quad + O\left(\left(\log \frac{K}{s_0 e^{rT}}\right)^{-\frac{1}{2}}\right) \\
&= \frac{\sqrt{2}}{\sqrt{T}} \sqrt{-\alpha - 1} \sqrt{\log \frac{K}{s_0 e^{rT}}} \sqrt{1 + s_1(K)} \\
&\quad - \frac{\sqrt{2}}{\sqrt{T}} \sqrt{-\alpha - 2} \sqrt{\log \frac{K}{s_0 e^{rT}}} \sqrt{1 + s_2(K)} \\
&\quad + O\left(\left(\log \frac{K}{s_0 e^{rT}}\right)^{-\frac{1}{2}}\right)
\end{aligned} \tag{48}$$

as $K \rightarrow \infty$, where

$$s_1(K) = \frac{\log \frac{1}{h(K)} - \frac{1}{2} \log \log \frac{K}{s_0 e^{rT}}}{(-\alpha - 1) \log \frac{K}{s_0 e^{rT}}}$$

and

$$s_2(K) = \frac{\log \frac{1}{h(K)} - \frac{1}{2} \log \log \frac{K}{s_0 e^{rT}}}{(-\alpha - 2) \log \frac{K}{s_0 e^{rT}}}.$$

Next, using the formula $\sqrt{1+s} = 1 + \frac{1}{2}s + O(s^2)$ as $s \rightarrow 0$ in (48), we obtain

$$\begin{aligned}
I(K) &= \frac{\sqrt{2}}{\sqrt{T}} \left[\sqrt{-\alpha-1} - \sqrt{-\alpha-2} \right] \sqrt{\log \frac{K}{s_0 e^{rT}}} \\
&\quad + \frac{1}{\sqrt{2T}} \left[\frac{1}{\sqrt{-\alpha-1}} - \frac{1}{\sqrt{-\alpha-2}} \right] \frac{\log \frac{1}{h(K)} - \frac{1}{2} \log \log \frac{K}{s_0 e^{rT}}}{\sqrt{\log \frac{K}{s_0 e^{rT}}}} \\
&\quad + O \left(\left(\log \frac{K}{s_0 e^{rT}} \right)^{-\frac{1}{2}} \right) \\
&= \frac{\sqrt{2}}{\sqrt{T}} \left[\sqrt{-\alpha-1} - \sqrt{-\alpha-2} \right] \sqrt{\log \frac{K}{s_0 e^{rT}}} \\
&\quad + \left[\frac{1}{\sqrt{-\alpha-2}} - \frac{1}{\sqrt{-\alpha-1}} \right] \frac{\tilde{B}}{T^{\frac{3}{4}} (8\tilde{C} + T)^{\frac{1}{4}}} \\
&\quad + \frac{1-n_1}{4} \frac{\log \log \frac{K}{s_0 e^{rT}}}{\sqrt{\log \frac{K}{s_0 e^{rT}}}} + O \left(\left(\log \frac{K}{s_0 e^{rT}} \right)^{-\frac{1}{2}} \right)
\end{aligned}$$

as $K \rightarrow \infty$. Finally, taking into account the definition of the parameter α , we see that (46) holds.

The proof of Theorem 4 is thus completed. \square

In terms of the log-moneyness $k = \log \frac{K}{s_0 e^{rT}}$, Theorem 4 can be formulated as follows.

Theorem 6. *The following formula holds for the implied volatility $k \mapsto I(k)$ as $k \rightarrow \infty$:*

$$\begin{aligned}
I(k) &= \frac{1}{T^{\frac{3}{4}}} \left[\left(\sqrt{8\tilde{C} + T} + \sqrt{T} \right)^{\frac{1}{2}} - \left(\sqrt{8\tilde{C} + T} - \sqrt{T} \right)^{\frac{1}{2}} \right] \sqrt{k} \\
&\quad + \frac{\sqrt{2}\tilde{B}}{T^{\frac{1}{2}} (8\tilde{C} + T)^{\frac{1}{4}}} \left[\frac{1}{\left(\sqrt{8\tilde{C} + T} - \sqrt{T} \right)^{\frac{1}{2}}} - \frac{1}{\left(\sqrt{8\tilde{C} + T} + \sqrt{T} \right)^{\frac{1}{2}}} \right] \\
&\quad + \frac{1-n_1}{4} \frac{\log k}{\sqrt{k}} + O \left(k^{-\frac{1}{2}} \right). \tag{49}
\end{aligned}$$

Note that if $n_1 = 1$, then the third term in the asymptotic expansion of the implied volatility in formula (46) vanishes. On the other hand, if the Gaussian process X is centered, then $\tilde{B} = 0$, and the second term on the tight-hand side of (46) vanishes. The next corollary takes those facts into account.

Corollary 7. *The following are true:*

(i) *If $n_1 = 1$, then as $K \rightarrow \infty$,*

$$\begin{aligned} I(K) &= \frac{2\lambda_1^{\frac{1}{4}}}{T^{\frac{1}{2}} \left[(\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{1}{2}} + (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{1}{2}} \right]} \sqrt{\log \frac{K}{s_0 e^{rT}}} \\ &\quad + \frac{\sqrt{2}\delta}{T^{\frac{1}{2}}(4 + \lambda_1)^{\frac{1}{4}} \left[(\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{1}{2}} + (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{1}{2}} \right]} \\ &\quad + O\left(\left(\log \frac{K}{s_0 e^{rT}}\right)^{-\frac{1}{2}}\right). \end{aligned}$$

(ii) *If X is a centered Gaussian process, then as $K \rightarrow \infty$,*

$$\begin{aligned} I(K) &= \frac{2\lambda_1^{\frac{1}{4}}}{T^{\frac{1}{2}} \left[(\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{1}{2}} + (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{1}{2}} \right]} \sqrt{\log \frac{K}{s_0 e^{rT}}} \\ &\quad + \frac{1 - n_1}{4} \frac{\log \log \frac{K}{s_0 e^{rT}}}{\sqrt{\log \frac{K}{s_0 e^{rT}}}} + O\left(\left(\log \frac{K}{s_0 e^{rT}}\right)^{-\frac{1}{2}}\right). \end{aligned}$$

(iii) *If X is a centered Gaussian process and $n_1 = 1$, then as $K \rightarrow \infty$,*

$$\begin{aligned} I(K) &= \frac{2\lambda_1^{\frac{1}{4}}}{T^{\frac{1}{2}} \left[(\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{1}{2}} + (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{1}{2}} \right]} \sqrt{\log \frac{K}{s_0 e^{rT}}} \\ &\quad + O\left(\left(\log \frac{K}{s_0 e^{rT}}\right)^{-\frac{1}{2}}\right). \end{aligned}$$

The constants in Corollary 7 are obtained from those in Theorem 4, by taking into account (47). These constants depend only on the following characteristics of the process X : The largest eigenvalue λ_1 of the covariance operator \mathcal{K} on $[0, T]$, and the Fourier coefficient δ_1 of the mean function m with respect to the corresponding eigenfunction e_1 . We refer to Section 1.3 for further discussion.

Since the processes X and W in (35) are independent, the stochastic volatility model described in (35) belongs to the class of the so-called symmetric models (see Section 9.8 in [21]). It is known that for symmetric models,

$$I(K) = I\left(\frac{(s_0 e^{rT})^2}{K}\right) \quad \text{for all } K > 0. \quad (50)$$

The next statements can be established using Theorem 4, Corollary 7, and formula (50).

Theorem 8. *The following formula holds for the function $K \mapsto I(K)$ as $K \rightarrow 0$:*

$$\begin{aligned}
I(K) = & \frac{1}{T^{\frac{3}{4}}} \left[\left(\sqrt{8\tilde{C} + T} + \sqrt{T} \right)^{\frac{1}{2}} - \left(\sqrt{8\tilde{C} + T} - \sqrt{T} \right)^{\frac{1}{2}} \right] \sqrt{\log \frac{s_0 e^{rT}}{K}} \\
& + \frac{\sqrt{2\tilde{B}}}{T^{\frac{1}{2}}(8\tilde{C} + T)^{\frac{1}{4}}} \left[\frac{1}{\left(\sqrt{8\tilde{C} + T} - \sqrt{T} \right)^{\frac{1}{2}}} - \frac{1}{\left(\sqrt{8\tilde{C} + T} + \sqrt{T} \right)^{\frac{1}{2}}} \right] \\
& + \frac{1 - n_1}{4} \frac{\log \log \frac{s_0 e^{rT}}{K}}{\sqrt{\log \frac{s_0 e^{rT}}{K}}} + O \left(\left(\log \frac{s_0 e^{rT}}{K} \right)^{-\frac{1}{2}} \right). \tag{51}
\end{aligned}$$

The constants in (51) are the same as in Theorem 4.

Corollary 9. *The following are true:*

(i) *If $n_1 = 1$, then as $K \rightarrow 0$,*

$$\begin{aligned}
I(K) = & \frac{2\lambda_1^{\frac{1}{4}}}{T^{\frac{1}{2}} \left[(\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{1}{2}} + (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{1}{2}} \right]} \sqrt{\log \frac{s_0 e^{rT}}{K}} \\
& + \frac{\sqrt{2}\delta_1}{T^{\frac{1}{2}}(4 + \lambda_1)^{\frac{1}{4}} \left[(\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{1}{2}} + (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{1}{2}} \right]} \\
& + O \left(\left(\log \frac{s_0 e^{rT}}{K} \right)^{-\frac{1}{2}} \right). \tag{52}
\end{aligned}$$

(ii) *If X is a centered Gaussian process, then as $K \rightarrow 0$,*

$$\begin{aligned}
I(K) = & \frac{2\lambda_1^{\frac{1}{4}}}{T^{\frac{1}{2}} \left[(\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{1}{2}} + (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{1}{2}} \right]} \sqrt{\log \frac{s_0 e^{rT}}{K}} \\
& + \frac{1 - n_1}{4} \frac{\log \log \frac{s_0 e^{rT}}{K}}{\sqrt{\log \frac{s_0 e^{rT}}{K}}} + O \left(\left(\log \frac{s_0 e^{rT}}{K} \right)^{-\frac{1}{2}} \right).
\end{aligned}$$

(iii) *If X is a centered Gaussian process and $n_1 = 1$, then as $K \rightarrow 0$,*

$$I(K) = \frac{2\lambda_1^{\frac{1}{4}}}{T^{\frac{1}{2}} \left[(\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{1}{2}} + (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{1}{2}} \right]} \sqrt{\log \frac{s_0 e^{rT}}{K}} + O \left(\left(\log \frac{s_0 e^{rT}}{K} \right)^{-\frac{1}{2}} \right).$$

6. IMPLIED VOLATILITY IN THE UNCORRELATED STEIN-STEIN MODEL

The classical Stein-Stein model is an important special example of a Gaussian stochastic volatility model. The Stein-Stein model was introduced in [33]. The volatility in the uncorrelated Stein-Stein model is the absolute value of an Ornstein-Uhlenbeck process

with a constant initial condition m_0 . In this section, we also consider a generalization of the Stein-Stein model, in which the initial condition for the volatility process is a random variable X_0 . Of our interest in the present section is a Gaussian stochastic volatility model with the process X satisfying the equation $dX_t = q(m - X_t)dt + \sigma dZ_t$. Here $q > 0$, $m \geq 0$, and $\sigma > 0$. It will be assumed that the initial condition X_0 is a Gaussian random variable with mean m_0 and variance σ_0^2 , independent of the process Z . It is known that

$$X_t = e^{-qt} X_0 + (1 - e^{-qt})m + \sigma e^{-qt} \int_0^t e^{qu} dZ_u, \quad t \geq 0. \quad (53)$$

If $\sigma_0 = 0$, then the initial condition is equal to the constant m_0 . The mean function of the process X is given by

$$m(t) = e^{-qt} m_0 + (1 - e^{-qt})m, \quad (54)$$

and its covariance function is as follows:

$$Q(t, s) = e^{-q(t+s)} \left\{ \sigma_0^2 + \frac{\sigma^2}{2q} \left(e^{2q \min(t,s)} - 1 \right) \right\}.$$

Therefore, the following formula holds for the variance function:

$$\sigma_t^2 = \frac{\sigma^2}{2q} + e^{-2qt} \left(\sigma_0^2 - \frac{\sigma^2}{2q} \right),$$

and hence, if $\sigma_0^2 = \frac{\sigma^2}{2q}$, then the process $X_t - m(t)$, $t \in [0, T]$, is centered and stationary. In this case, the covariance function is given by

$$Q(t, s) = \frac{\sigma^2}{2q} e^{-q|t-s|}.$$

The Karhunen-Loève expansion of the Ornstein-Uhlenbeck process is known explicitly (see [11]). Denote by w_n the increasingly sorted sequence of the positive solutions to the equation

$$\sigma^2 w \cos(wT) + (q\sigma^2 - w^2\sigma_0^2 - q^2\sigma_0^2) \sin(wT) = 0. \quad (55)$$

If $\sigma_0 = 0$, then the equation in (55) becomes

$$w \cos(wT) + q \sin(wT) = 0. \quad (56)$$

For the OU process in (53) with $\sigma_0 \neq 0$, we have $n_k = 1$ for all $k \geq 1$;

$$\lambda_n = \frac{\sigma^2}{w_n^2 + q^2} \quad (57)$$

for all $n \geq 1$; and

$$e_n(t) = K_n [\sigma_0^2 w_n \cos(w_n t) + (\sigma^2 - q\sigma_0^2) \sin(w_n t)] \quad (58)$$

for all $n \geq 1$ and $t \in [0, T]$. The constant K_n in (58) is determined from

$$\begin{aligned} \frac{1}{K_n^2} &= \frac{1}{2w_n} \sigma_0^2 (\sigma^2 - q\sigma_0^2) (1 - \cos(2w_n T)) + \frac{1}{2} \sigma_0^4 w_n^2 \left(T + \frac{1}{2w_n} \sin(2w_n T) \right) \\ &\quad + \frac{1}{2} (\sigma^2 - q\sigma_0^2)^2 \left(T - \frac{1}{2w_n} \sin(2w_n T) \right) \end{aligned} \quad (59)$$

for all $n \geq 1$. On the other hand, if $\sigma_0 = 0$, then λ_n is given by (57), while the functions e_n are defined by

$$e_n(t) = \frac{1}{\sqrt{\frac{T}{2} - \frac{\sin(2w_n T)}{4w_n}}} \sin(w_n t) \quad (60)$$

for all $n \geq 1$ and $t \in [0, T]$.

By the Karhunen-Loève theorem, the Ornstein-Uhlenbeck process X in (53) can be represented as follows:

$$X_t = e^{-qt} m_0 + (1 - e^{-qt}) m + \sum_{n=1}^{\infty} \sqrt{\lambda_n} e_n(t) Z_n$$

where $\{Z_n\}_{n \geq 1}$ is an i.i.d. sequence of standard normal variables. The eigenvalues λ_n , $n \geq 1$, and the eigenfunctions e_n , $n \geq 1$, are given by (57) and (58) if $\sigma_0 \neq 0$, and by (57) and (60) if $\sigma_0 = 0$. Recall that the numbers w_n , $n \geq 1$, in (57) are solutions to the equation in (55) if $\sigma_0 \neq 0$, and to the equation in (56) if $\sigma_0 = 0$. We refer the interested reader to [11] for more details.

Our next goal is to discuss the constants in the asymptotic formulas for the implied volatility at extreme strikes in the Stein-Stein model (see the formulas in Corollaries 7 and 9). Since $n_1 = 1$ for any OU process, the only parameters needed to compute the above-mentioned constants are λ_1 and δ_1 . If $\sigma_0 \neq 0$, then we have

$$\lambda_1 = \frac{\sigma^2}{w_1^2 + q^2}, \quad (61)$$

where w_1 is the smallest strictly positive solution to the equation in (55).

The next assertion provides explicit formulas for the number $\delta_1 = \int_0^T m(t) e_1(t) dt$.

Lemma 10. (i) For the generalized uncorrelated Stein-Stein model with $\sigma_0 \neq 0$,

$$\begin{aligned} \delta_1 = & \frac{K_1 m (\sigma^2 - q \sigma_0^2) (1 - \cos(w_1 T))}{w_1} + K_1 \sigma_0^2 \sin(w_1 T) [(m_0 - m) e^{-qT} + m] \\ & + K_1 \sigma^2 (m_0 - m) \frac{w_1 [1 - e^{-qT} \cos(w_1 T)] - q e^{-qT} \sin(w_1 T)}{q^2 + w_1^2}, \end{aligned} \quad (62)$$

where the constant K_1 is determined from (59) with $n = 1$. The symbol w_1 in (62) stands for the smallest strictly positive solution to (55).

(ii) For the uncorrelated Stein-Stein model with $X_0 = m_0$ \mathbb{P} -almost surely,

$$\delta_1 = \frac{mq^2(1 - \cos(w_1 T)) + w_1^2(m_0 - m \cos(w_1 T))}{w_1(q^2 + w_1^2) \sqrt{\frac{T}{2} - \frac{\sin(2w_1 T)}{4w_1}}}. \quad (63)$$

Proof. Taking into account (54) and (58), we see that

$$\begin{aligned} \delta_1 = & b_1 \int_0^T \cos(w_1 t) dt + b_2 \int_0^T e^{-qt} \cos(w_1 t) dt \\ & + b_3 \int_0^T \sin(w_1 t) dt + b_4 \int_0^T e^{-qt} \sin(w_1 t) dt, \end{aligned} \quad (64)$$

where

$$\begin{aligned} b_1 &= mK_1\sigma_0^2w_1, \quad b_2 = (m_0 - m)K_1\sigma_0^2w_1, \\ b_3 &= mK_1(\sigma^2 - q\sigma_0^2), \quad \text{and} \quad b_4 = (m_0 - m)K_1(\sigma^2 - q\sigma_0^2). \end{aligned} \quad (65)$$

It remains to evaluate the integrals in (64). We have

$$\int_0^T \cos(w_1 t) dt = \frac{\sin(w_1 T)}{w_1}, \quad (66)$$

$$\int_0^T e^{-qt} \cos(w_1 t) dt = \frac{q[1 - e^{-qT} \cos(w_1 T)] + w_1 e^{-qT} \sin(w_1 T)}{q^2 + w_1^2}, \quad (67)$$

$$\int_0^T \sin(w_1 t) dt = \frac{1 - \cos(w_1 T)}{w_1}, \quad (68)$$

and

$$\int_0^T e^{-qt} \sin(w_1 t) dt = \frac{w_1[1 - e^{-qT} \cos(w_1 T)] - qe^{-qT} \sin(w_1 T)}{q^2 + w_1^2}. \quad (69)$$

In the proof of (67) and (69), we use the integration by parts formula twice. Now, taking into account formulas (64-69) and making simplifications, we establish formula (62).

Next, suppose $\sigma_0 = 0$. Then (62) implies that

$$\begin{aligned} \delta_1 &= \frac{m}{\sqrt{\frac{T}{2} - \frac{\sin(2w_1 T)}{4w_1}}} \frac{1 - \cos(w_1 T)}{w_1} \\ &\quad + \frac{m_0 - m}{\sqrt{\frac{T}{2} - \frac{\sin(2w_1 T)}{4w_1}}} \frac{w_1[1 - e^{-qT} \cos(w_1 T)] - qe^{-qT} \sin(w_1 T)}{q^2 + w_1^2}, \end{aligned}$$

where w_1 denotes the smallest strictly positive solution to (56). It is not hard to see, using the equality $w_1 \cos(w_1 T) + q \sin(w_1 T) = 0$, that

$$\delta_1 = \frac{m}{\sqrt{\frac{T}{2} - \frac{\sin(2w_1 T)}{4w_1}}} \frac{1 - \cos(w_1 T)}{w_1} + \frac{m_0 - m}{\sqrt{\frac{T}{2} - \frac{\sin(2w_1 T)}{4w_1}}} \frac{w_1}{q^2 + w_1^2}, \quad (70)$$

and it is clear that (70) and (63) are equivalent.

This completes the proof of Lemma 10. \square

Remark 11. Since for the generalized Stein-Stein model with a random initial condition we have $n_1 = 1$, one can use the asymptotic formulas in the first parts of Corollaries 7 and 9 to characterize the wing-behavior of the implied volatility. The dependence of the parameters λ_1 and δ_1 , appearing in those formulas, on the model parameters is described in (61), (62), and (63). Originally, sharp asymptotic formulas for the implied volatility at extreme strikes in the uncorrelated Stein-Stein model with $X_0 = m_0$ were obtained in [24] (see also Section 10.5 in [21]). However, explicit expressions, obtained in [24, 21] for the coefficients in the asymptotic formulas for the implied volatility in the Stein-Stein model, are significantly more complicated than those found in the present paper.

7. NUMERICAL ILLUSTRATION

A basic calibration strategy when presented with asymptotic results such as those given in this paper is to assume one can place oneself in the corresponding regime, and then determine model parameters by reading asymptotic coefficient off of market option prices. We now illustrate how this strategy can produce positive results, and discuss its limitations, in the context of our Corollary 9 part (i), i.e. when the top of the KL spectrum is simple ($n_1 = 1$). In other words, the idea is to ignore the big O term in the asymptotic (52), and calibrate parameters to the remaining coefficients. Denoting the discounted log-moneyness $\log(S_0 e^{rT}/K)$ by k for compactness of notation, we thus have, for $|k|$ sufficiently large,

$$I(k) \simeq L\sqrt{|k|} + M, \quad (71)$$

for two constants L and M , which can, in principle, be read off of market data. By the explicit expressions for these constants in (52) in terms of λ_1 and δ_1 , we then express the latter in terms of L and M as

$$\begin{aligned} \lambda_1 &= \frac{64T^2L^4}{(4 - T^2L^4)^2}, \\ \delta_1 &= \frac{4\sqrt{2TM}\sqrt{4 + T^2L^4}}{4 - T^2L^4}. \end{aligned} \quad (72)$$

Here we use (71). One notices that, conveniently, λ_1 can be calibrated using only the coefficient L , while given L , δ_1 is then proportional to M .

At this stage, one may simply conclude that the extreme strike asymptotics given in the market are consistent with any Gaussian volatility model whose top of eigenstructure is represented by the values computed in the above expressions for λ_1 and δ_1 . However, practitioners will prefer to determine a more specific model, perhaps by choosing a classical parametric one, and using other non-asymptotic-calibration techniques for estimating some of its parameters. The expressions in (72) can then be used to pin down other parameters by calibration, as long as one can relate the model's parameters to its KL pair (λ_1, δ_1) , whether analytically or numerically.

We provide illustrations of this strategy in two cases: the stationary Stein-Stein model, where the KL expansion is known semi-explicitly, and the Stein-Stein model's long-memory version, where the volatility is also known as the fractional Ornstein-Uhlenbeck (fOU) model, and the KL expansion is computed numerically. The data we use is also generated numerically: for each model, we compute option prices and their corresponding implied volatilities, by classical Monte-Carlo, given that the underlying pair of stochastic processes is readily simulated. Specifically, in the Stein-Stein (standard OU) case, 10^6 paths are generated via Euler's method based on discretizing the stochastic differential equation satisfied by X started from a r.v. sampled from X 's stationary distribution, and the explicit expression for $\log S$ given X , also approximated via Euler with the same time steps; 10^3 time steps are used in $[0, T]$. Option prices are derived by computing average payoffs over the 10^6 paths. The details are well known, and are omitted. In the fOU case, the exact same methodology is used, except that one must specify the technique used to simulate increments of the fBm process which drives X : we used the circulant method,

which is based on fBm's spectral properties, and was proposed by A.T. Dieker in a 2002 thesis: see [15, 16].

Given this simulated data, before embarking on the task of calibrating parameters, to ensure that our methodology is relevant in practice, it is important to discuss liquidity issues. It is known that the out-of-the-money call options market is poorly liquid, implying that the large strike asymptotics for call and IV prices are typically not visible in the data. We concentrate instead on small strike asymptotics. There, depending on the market segment, options with maturities on the scale of one or several months can be liquid with small bid-ask spread for log moneyness as far down as -1 . Convincing visual evidence of this can be found in Figures 3 and 4 in [20] which report 2011 data for SPX options. Based on this, we will use the value -1 as an illustrative lower bound for k : beyond this lower bound, liquidity is insufficient to measure IV. This immediately creates a problem due to the similar magnitude, in this range of k , of the constant term M and the next term in the expansion whose order we can only guarantee to be no larger than $1/\sqrt{-k}$, according to the big O term in Corollary 9 part (i).

In spite of this, calibration of λ_1 to L as in (72) works well in practice, as our examples below show. We also find that the inability to neglect $O(1/\sqrt{-k})$ compared to M introduces a bias in the calibration, whose magnitude, while not constant, gives a good idea of the relative size of the next term in the small strike expansion, motivating the open problem of looking for this next term, and perhaps the following one.

We begin with the stationary uncorrelated Stein-Stein model with constant mean-reversion level m , rate of mean reversion q , and so called vol-vol parameter σ . The systems of equations needed to perform calibration here have a somewhat triangular structure. According to Section 6, if one is to calibrate q , access to δ_1 is needed, if one may rely on independent knowledge of the level of mean reversion m . Specifically, one would solve the following system

$$\begin{aligned} q \sin(wT) + w \cos(wT) &= 0 \\ C_1 \left(\sin(wT) + \frac{q}{w} (1 - \cos(wT)) \right) &= \frac{\delta_1}{m} \end{aligned}$$

where unfortunately the constant C_1 also depends on the unknowns (q, w) in the following non-trivial way:

$$\frac{1}{C_1^2} = \frac{q}{2} (1 - \cos(2wT)) + \frac{w^2}{2} \left(T + \frac{\sin(2wT)}{2w} \right) + \frac{q^2}{2} \left(T - \frac{\sin(2wT)}{2w} \right).$$

Since q is not fixed, the task of determining which value of w represents the minimal solution of the first equation above, given the large number of solutions to the above system, is difficult. We did not pursue this avenue further for this reason, and because the aforementioned bias due to liquidity means that one would not be in the right asymptotic regime.

Nevertheless, in the Stein-Stein model, our asymptotics allow us to calibrate the vol-vol parameter σ . The equation for finding σ given a measurement of L and prior knowledge of q is much simpler. Indeed, since q is assumed given, the base frequency w is computed easily as the smallest positive solution of (56). Independently of this, based on L , the first

relation in (72) determine λ_1 . Then according to equation (61), we obtain immediately

$$\sigma^2 = \lambda_1 (w_1^2 + q^2).$$

We simulated IV data for the call option with $m = 0.2$ (signifying a typical mean level of volatility of 20%), $q = 7$ (fast mean reversion, every eight weeks or so), and $\sigma = 1.2$ (high level of volatility uncertainty). How to estimate L from the data is not unambiguous. We adopt a least-squares method, on an interval of k -values of fixed length; after experimentation, as a rule of thumb, an interval of length 0.15 provides a good balance between providing a local estimate and drawing on enough datapoints. We start the interval as far to the left as possible while avoiding any range where the data representing the function $k \mapsto I(k)$ may exhibit clear signs of convexity. This is in an effort to coincide with values of k for which sufficient liquidity exists in practice. As a guide to assess this liquidity, we use the study reported in [20, Section 5.4], which we mentioned in the introduction. This liquidity depends heavily on the option maturity.

For a one-month ($T = 1/12$) option, the above calibration method reports $\sigma = 1.183$, based on the interval $k \in [-0.85, -0.70]$. For a three-month option ($T = 1/4$), the method reports $\sigma = 1.193$, based on the interval $[-1.25, -1.10]$. Other calibrations, not reported here because of their similarity with these two, show that calibration accuracy increases with more liquid options, which is consistent with the intuition that being able to use intervals further to the left should allow a better match with the asymptotic regime (71). The graphs of the data versus the asymptotic curve, showing excellent agreement for the first term $L\sqrt{-k}$, are given in Fig. 1 and Fig. 2.

As mentioned above, this agreement explains why this calibration for σ works well even though there is a clear bias between the curve $k \mapsto L\sqrt{-k} + M$ (solid lines in Fig. 1 and Fig. 2) and the true (simulated) IV data (blue data points in Fig. 1 and Fig. 2). One may immediately conjecture that the gap between the solid and discrete lines is representative of the term $O(1/\sqrt{-k})$ in Corollary 9, in the sense that the next term in the expansion might be of the form $N/\sqrt{-k}$ for some constant N , and that our simulation might give us access to the value of N . This does not appear to be quite the case. When graphing the expression for $[I(k) - (L\sqrt{-k} + M)]\sqrt{-k}$ using data from the previous figure, instead of finding a quantity which is largely constant, we find values which vary non-linearly between 0.02 and 0.026, as seen in Fig. 3. This is indicative that, to have access to calibration based on the value of M , it might be necessary to find more terms in the asymptotic expansion.

We now explain how to include one more term in formula (3), specifically, how to obtain the following expansion:

$$I(k) = L\sqrt{-k} + M + \frac{N}{\sqrt{-k}} + O\left(-\frac{1}{k}\right) \quad (73)$$

as $k \rightarrow -\infty$ when $n_1 = 1$. It is possible to establish formula (73) by using formula (41) instead of a weaker two-sided estimate given in (43), and also formula (39) and Theorem 5.6 in [21], characterizing the asymptotic behavior of log-normal integrals. Combining these results, we obtain a sharp asymptotic formula for the density D_t with a relative error

estimate $O\left((\log x)^{-\frac{1}{2}}\right)$ as $x \rightarrow \infty$. After that, Theorem 14 in [25] can be used to establish (73). Obtaining the expression for the constant N is tractable but sufficiently non-trivial that we omit the details. Even in the case where the volatility process is centered and $n_1 = 1$, this N depends on all λ_k with $k \geq 1$ and n_k with $k \geq 2$. We also conjecture that when $n_1 = 1$ and $n_2 = 1$,

$$I(k) = L\sqrt{-k} + M + \frac{N}{\sqrt{-k}} + \frac{P}{-k} + O\left(\frac{1}{(-k)^{3/2}}\right) \quad (74)$$

as $k \rightarrow -\infty$. This conjecture would not be sufficient to guarantee that including the four terms identified above would be enough to provide a good approximation when $|k|$ is restricted to values around 1 or lower, since there a series in negative powers of $\sqrt{-k}$ will diverge unless its coefficients decay fast enough. Yet Fig. 1 and Fig. 3 show, at least in those particular examples, that the discrepancy which might correspond to $N/\sqrt{-k} + P/(-k)$ is not large, and that the term $P/(-k)$ would only be needed to account for a moderate uncertainty level on the value of N , ranging from 0.02 to 0.026. This gives hope that the first three or four terms in the expansion (74) could provide an excellent approximation to the IV even for values of $|k|$ lower than 1.

While the conjecture in (74) may be correct, the last analysis above attempts to push the boundaries of its possible validity beyond its natural range because of option market liquidity considerations. This empirical consideration seems to rely on the class of Gaussian volatility models one chooses. This becomes evident in the second class of examples we present, that of fOU volatility. There, as we will see, the bias between $I(k)$ and $L\sqrt{-k} + M$ cannot reasonably be considered small, so that the conjecture in (74) may not be helpful. Nevertheless, our calibration strategy still produces good results when applied to the fOU's memory parameter. We now provide the details.

The continuous-time fOU volatility model was introduced in [8] as a way to model long-range dependence in non-linear functionals of stock returns, while preserving the uncorrelated semi-martingale structure at the level of returns themselves. This is the model for X in (2) where the process Z is a fractional Brownian motion, i.e. the continuous Gaussian process started at 0 with covariance determined by $\mathbb{E}\left[(Z_t - Z_s)^2\right] = |t - s|^{2H}$, with "Hurst" parameter $H \in (0.5, 1)$. In [7], it was shown empirically that standard statistical methods for long-memory data are inadequate for estimating H . This difficulty can be attributed to the fact that the volatility process X can have non-stationary increments. In addition, some of the classical methods use path regularity or self-similarity as a proxy for long memory, which cannot be exploited in practice since there is a lower limit to how frequently observations can be made without running into microstructure noise. To make matters worse, the process X is not directly observed; in such a partial observation case, a general theoretical result was given in [32], by which the rate of convergence of any estimator of H cannot exceed an optimal H -dependent rate which is always slower than $N^{-1/4}$, where N is the number of observations. Given the non-stationarity of the parameter H on a monthly scale, a realistic time series at the highest observation frequency where microstructure noise can be ingored (e.g. one stock observation every 5 minutes) would not permit even the optimal estimators described in [32] from pinning down a

value of H with any acceptable confidence level. The work in [7] proposes a calibration technique based on a straightforward comparison of simulated and market option prices to determine H . Our strategy herein is similar, but based on implied volatility.

We consider the fOU model described above with the following parameters: $T = 1/12$, $m = 0.2$, $q = 7$, $\sigma = 1.2$, with different values of the Hurst parameter $H \in \{0.51, 0.55, 0.60, 0.70, 0.80\}$. As mentioned above, we simulated option prices using standard Monte Carlo, where the fOU process is produced by A.T. Dieker's circulant method. Since the values of λ_1 for each $H > 0.5$ are not known explicitly or semi-explicitly, we resorted to the method developed in by S. Corlay in [9] for optimal quantification: there, the infinite-dimensional eigenvalue problem is converted to a matrix eigenvalue problem which uses a low-order quadrature rule for approximating integrals (a trapezoidal rule is recommended), after which a Richardson-Romberg extrapolation is used to improve accuracy. We repeat this procedure for the fOU process with the above parameters, for each value of H from 0.50 to 0.99, with increments of 0.01. The corresponding values we obtain for λ_1 in each case are collected in the following table:

$H =$	0.50	0.51	0.52	0.53	0.54	0.55	0.56	0.57	0.58	0.59
$\lambda_1 =$	0.00713	0.00697	0.00682	0.00667	0.00653	0.00638	0.00624	0.00610	0.00597	0.00584
$H =$	0.60	0.61	0.62	0.63	0.64	0.65	0.66	0.67	0.68	0.69
$\lambda_1 =$	0.00571	0.00558	0.00546	0.00534	0.00522	0.00511	0.00499	0.00488	0.00478	0.00467
$H =$	0.70	0.71	0.72	0.73	0.74	0.75	0.76	0.77	0.78	0.79
$\lambda_1 =$	0.00457	0.00447	0.00437	0.00428	0.00419	0.00410	0.00401	0.00392	0.00384	0.00376
$H =$	0.80	0.81	0.82	0.83	0.84	0.85	0.86	0.87	0.88	0.89
$\lambda_1 =$	0.00368	0.00360	0.00353	0.00345	0.00338	0.00331	0.00324	0.00317	0.00311	0.00305
$H =$	0.90	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99
$\lambda_1 =$	0.00299	0.00293	0.00287	0.00281	0.00275	0.00270	0.00265	0.00260	0.00255	0.00250

Our illustration of the calibration method then consists of starting with simulated IV data for a fOU model with fixed H , then implementing the same procedure as in the Stein-Stein example to determine the value of λ_1 by using an interval of length 0.15 which is realistic in terms of liquidity constraints, and then matching that value of λ_1 to the closest value in the above table, thereby concluding that the simulated data is consistent with the corresponding value of H in the table. The results of this method are summarized as follows

True H	0.51	0.55	0.60	0.70	0.80
Interval used	$[-1.0, -0.85]$	$[-1.0, -0.85]$	$[-1.0, -0.85]$	$[-1.0, -0.85]$	$[-1.0, -0.85]$
measured λ_1	$\lambda_1 = 0.0071$	$\lambda_1 = 0.0064$	$\lambda_1 = 0.0055$	$\lambda_1 = 0.0045$	$\lambda_1 = 0.0037$
calibrated H	0.50	0.55	0.62	0.71	0.80

Our method shows a good level of accuracy, though it does not appear capable of distinguishing between the case of no memory $H = 0.50$ and a simulation with very low memory ($H = 0.51$). The accuracy of the method has some sensitivity to the intervals being used, particularly for small true values of H : in that case the first term in the implied volatility asymptotic curve seems to dictate a need for including data points for log-moneyness k near -1 . For higher values of H , we noted greater robustness to the interval

being used to estimate L . For example, with a true $H = 0.70$ used in simulation, calibrated values of $H = 0.69$ and $H = 0.70$ were obtained with the intervals $[-0.87, -0.72]$ and $[-0.90, -0.75]$ respectively.

More generally, one notes that the bias between the curve $L\sqrt{-k} + M$ and the simulated IV data is much larger when $H > 0.5$ than in the classical Stein-Stein case where $H = 0.5$, as illustrated in Fig. 4. This provides an even stronger call than in the case $H = 0.5$ to compute the constants N and P in the conjecture in relation (74), as functions of the first few KL eigen-elements.

We also computed the fOU's KL eigenvalues λ_2 and λ_3 for each $H \in \{0.50, 0.51, \dots, 0.99\}$, though these are not used in our calibration. They are nevertheless instructive since they indicate at what speed the KL expansions of X and of the integrated variance Y might converge. Without reporting all the values, we find that the ratio λ_2/λ_1 decreases from 0.11 for $H = 0.50$ to 0.0010 for $H = 0.99$, while for λ_3/λ_1 , the values are even smaller, decreasing from 0.033 to 0.00012 over the same range of H . If the values of N/L and P/L in our conjecture (74) are of the same order as the ratio λ_2/λ_1 , this could give additional hope that computing N, P in the fOU case might be enough to implement a two-parameter calibration, with only a few values of λ_i needed to approximate M, N, P .

Finally, by comparing Fig. 1 with the first graph in Fig. 4, where all OU parameters remain the same except for H which goes from 0.5 to 0.51, one notes that the difference $I(k) - (L\sqrt{-k} + M)$ jumps from values on the order of 0.02 to much larger values, on the order of 0.2. This indicates that the constant N in formula (73) might have a discontinuous dependence on H as H approaches 0.5; this phase transition between the case of no memory ($H = 0.5$) to that of long memory ($H > 0.5$) is presumably due to a jump in the value of KL elements as H approaches 0.5. This could lead to a method for determining whether long memory is present, but in the case of IV asymptotics it would be constrained by liquidity considerations. We will investigate all these questions in another article.

REFERENCES

- [1] R. J. Adler. *An introduction to continuity, extrema, and related topics for general Gaussian processes*. IMS Lecture Notes Monogr. Ser. **12**, 1990.
- [2] A. Alexanderian. A brief note on the Karhunen-Loève expansion, technical note, available on users.ices.utexas.edu/~alen/articles/KL.pdf
- [3] S. Benaïm and P. K. Friz. Regular variation and smile asymptotics. *Mathematical Finance* **19** (2009), 1-12.
- [4] S. Benaïm and P. K. Friz. Smile asymptotics, II: models with known moment generating function. *Journal of Applied Probability* **45** (2008), 16-32.
- [5] S. Benaïm, P. K. Friz, and R. Lee. On black-Scholes implied volatility at extreme strikes. In: R. Cont (Ed.), *Frontiers in Quantitative Finance: Volatility and Credit Risk Modeling*, Wiley, Hoboken (2009), 19-45.
- [6] R. Beran. The probabilities of noncentral quadratic forms. *The Annals of Statistics* **3** (1975), 969-974.
- [7] A. Chronopoulou and F. Viens. Stochastic volatility models with long-memory in discrete and continuous time. *Quantitative Finance*, **12** no. 4 (2012), 635-649.
- [8] F. Comte and E. Renault. Long Memory in Continuous-time Stochastic Volatility Models. *Mathematical Finance*, **8** (1998), 291-323.
- [9] S. Corlay. Quelques aspects de la quantification optimale, et applications en finance (in English, with French summary). Ph.D. Thesis. *Université Pierre et Marie Curie*, 2011.

- [10] S. Corlay. Properties of the Ornstein-Uhlenbeck bridge. Preprint, 2014, available on <https://hal.archives-ouvertes.fr/hal-00875342v4>
- [11] S. Corlay and G. Pagès. Functional quantization-based stratified sampling methods. Preprint, 2014, available on <https://hal.archives-ouvertes.fr/hal-00464088v3>
- [12] P. Deheuvels and G. Martynov. A Karhunen-Loeve decomposition of a Gaussian process generated by independent pairs of exponential random variables. *J. Funct. Anal.*, **255** (2008), 23263-2394.
- [13] J.-D. Deuschel, P. K. Friz, A. Jacquier and S. Violante. Marginal density expansions for diffusions and stochastic volatility I: Theoretical foundations. *Communications on Pure and Applied Mathematics* **67** (2014), 40-82.
- [14] J.-D. Deuschel, P. K. Friz, A. Jacquier and S. Violante. Marginal density expansions for diffusions and stochastic volatility I: Applications. *Communications on Pure and Applied Mathematics* **67** (2014), 321-350.
- [15] A. B. Dieker. Simulation of fractional Brownian motion. Master Thesis (2002), published by Universiteit Twente, 2004. Available at <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.114.8880>.
- [16] A. B. Dieker and M. Mandjes. On spectral simulation of fractional Brownian motion, *Probab. Engrg. Inform. Sci.*, **17** (2003), 417-434.
- [17] K. Gao and R. Lee. Asymptotics of Implied Volatility to Arbitrary Order. To appear in *Finance and Stochastics*, 2015. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1768383.
- [18] J. Gatheral. A parsimonious arbitrage-free implied volatility parametrization with application to the valuation of the volatility derivatives. In: *Global Derivatives and Risk Management*, Madrid, May 26, 2004.
- [19] J. Gatheral. *The Volatility Surface: A Practitioner's Guide*. Wiley New York 2006.
- [20] J. Gatheral and A. Jacquier. Arbitrage-free SVI volatility surfaces. *Quantitative Finance* **14** (1) (2014), 59-71.
- [21] A. Gulisashvili. *Analytically Tractable Stochastic Stock Price Models*, Springer-Verlag Berlin Heidelberg 2012.
- [22] A. Gulisashvili. Asymptotic formulas with error estimates for call pricing functions and the implied volatility at extreme strikes. *SIAM Journal on Financial Mathematics* **1** (2010), 609-641.
- [23] A. Gulisashvili. Asymptotic equivalence in Lee's moment formulas for the implied volatility, asset price models without moment explosions, and Piterbarg's conjecture. *International Journal of Theoretical and Applied Finance* **15** (2012), 1250020.
- [24] A. Gulisashvili and E. M. Stein. Asymptotic behavior of the stock price distribution density and implied volatility in stochastic volatility models. *Applied Mathematics and Optimization* **61** (2010), 287-315.
- [25] A. Gulisashvili and J. Vives. Asymptotic analysis of stock price densities and implied volatilities in mixed stochastic models. *SIAM J. Finan. Math.* **6** (2015), 158-188.
- [26] W. Hoeffding. On a theorem of V. M. Zolotarev. *Theor. Probab. Appl.* **9** (1964), 89-92.
- [27] I. A. Ibragimov and Y. A. Rozanov. *Gaussian Random Processes*, Springer-Verlag New York Heidelberg Berlin 1978.
- [28] R. Lee. The moment formula for implied volatility at extreme strikes. *Math. Finance* **14** (2004), 469-480.
- [29] I. Nourdin and G. Peccati. *Normal Approximations with Malliavin Calculus: From Stein's Method to Universality*. Cambridge U.P., 2012.
- [30] E. Renault and N. Touzi. Option Hedging and Implicit Volatilities, *Mathematical Finance* **6** (1996), 279-302.
- [31] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, Springer-Verlag Berlin 2004.
- [32] M. Rosenbaum. Estimation of the volatility persistence in a discretely observed diffusion model. *Stochastic Processes and their Applications*, **118** (8) (2008), 1434-1462.
- [33] E. Stein and J. Stein. Stock price distributions with stochastic volatility: an analytic approach. *Rev. Financ. Stud.* **4** (1991), 727-752.
- [34] A. M. Yaglom. *Correlation Theory of Stationary and Related Random functions, Vol. I*, Springer-Verlag New York 1987.
- [35] V. M. Zolotarev. Concerning a certain probability problem. *Theor. Probab. Appl.* **6** (1961), 201-204.

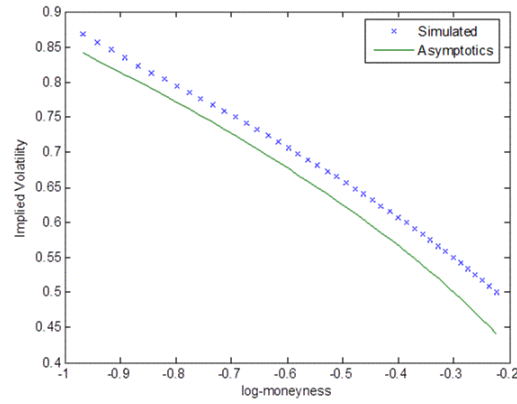


FIGURE 1. One-month IV for Stein-Stein model with parameters $m = 0.2$, $q = 7$, $\sigma = 1.2$

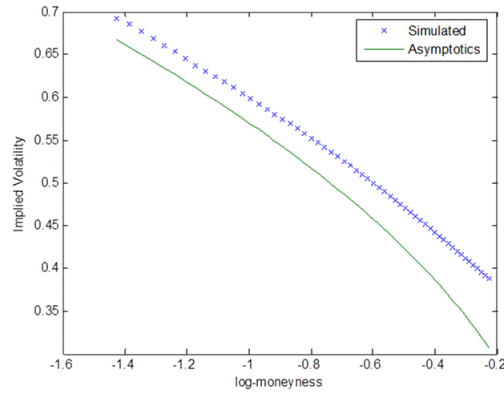


FIGURE 2. Three-month IV for Stein-Stein model with parameters $m = 0.2$, $q = 7$, $\sigma = 1.2$

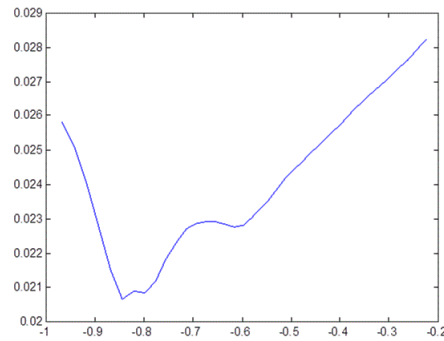


FIGURE 3. Graph of $\left[I(k) - \left(L\sqrt{-k} + M \right) \right] \sqrt{-k}$ with simulated data

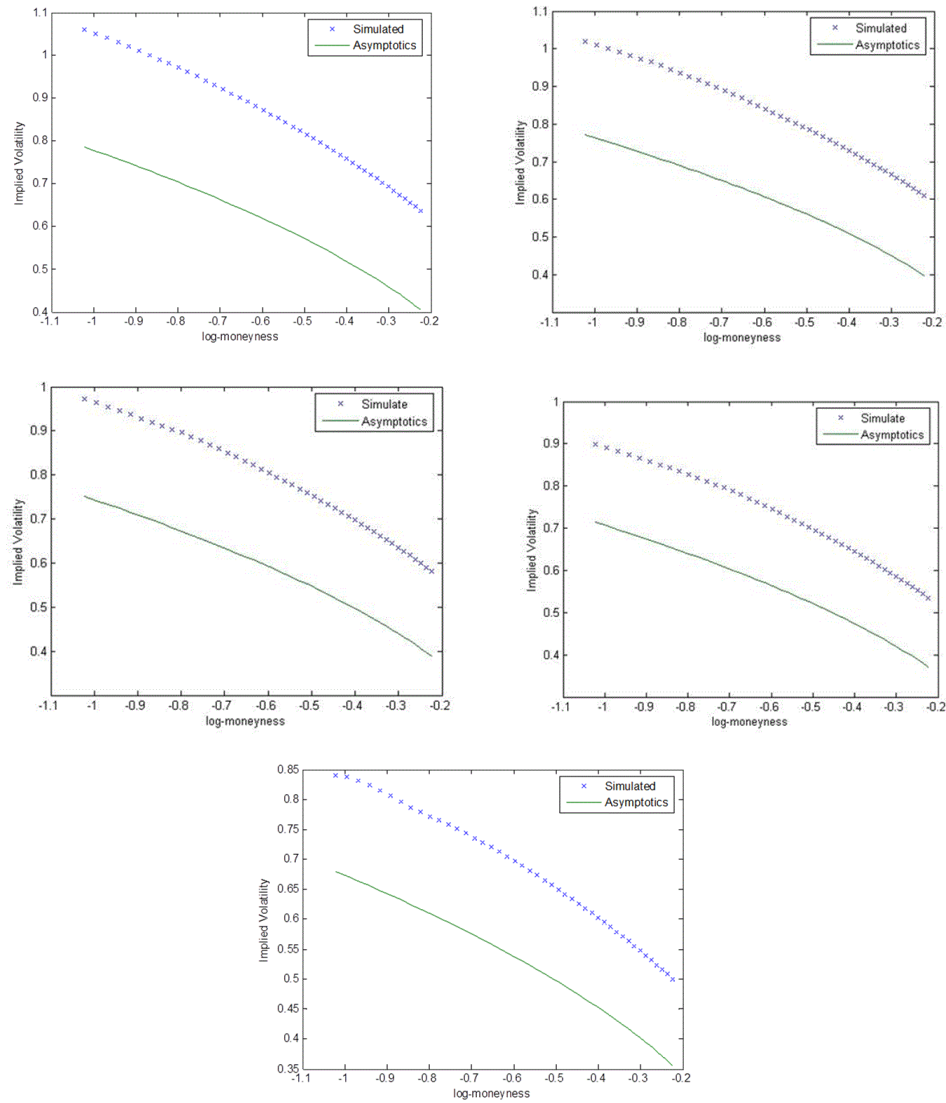


FIGURE 4. IV for fOU model with $H = 0.51$, $H = 0.55$, $H = 0.60$, $H = 0.70$, $H = 0.80$

DEPARTMENT OF MATHEMATICS, OHIO UNIVERSITY, ATHENS OH 45701, GULISASH@OHIO.EDU

DEPARTMENT OF STATISTICS, PURDUE UNIVERSITY, WEST LAFAYETTE, IN 47907, VIENS@PURDUE.EDU

DEPARTMENT OF MATHEMATICS, PURDUE UNIVERSITY, WEST LAFAYETTE, IN 47907, ZHANG407@MATH.PURDUE.EDU